

Automatic Detection of Myocontrol Failures Based upon Situational Context Information

Karoline Heiwolt¹, Claudio Zito², Markus Nowak³, Claudio Castellini³ and Rustam Stolkin²

Abstract—Myoelectric control systems for assistive devices are still unreliable. The user’s input signals can become unstable over time due to e.g. fatigue, electrode displacement, or sweat. Hence, such controllers need to be constantly updated and heavily rely on user feedback. In this paper, we present an automatic failure detection method which learns when plausible predictions become unreliable and model updates are necessary. Our key insight is to enhance the control system with a set of generative models that learn sensible behaviour for a desired task from human demonstration. We illustrate our approach on a grasping scenario in Virtual Reality, in which the user is asked to grasp a bottle on a table. From demonstration our model learns the reach-to-grasp motion from a resting position to two grasps (power grasp and tridigital grasp) and how to predict the most adequate grasp from local context, e.g. tridigital grasp on the bottle cap or around the bottleneck. By measuring the error between new grasp attempts and the model prediction, the system can effectively detect which input commands do not reflect the user’s intention. We evaluated our model in two cases: i) with both position and rotation information of the wrist pose, and ii) with only rotational information. Our results show that our approach detects statistically highly significant differences in error distributions with $p < 0.001$ between successful and failed grasp attempts in both cases.

I. INTRODUCTION

When designing a prosthesis, we aim to restore the full functionality of the missing body part, but as the functionality increases, it becomes more challenging to reliably control the device [1], [2], [3], [4], [5]. The demand for advanced and dexterous control systems among patients is high. And yet, the available myocontrol options are not very popular, with the main issue being reliability [1], [4], [6].

To overcome this obstacle, Gijsberts et al. [7] proposed a supervised incremental learning method, called *iRR-RFF*, that allows continuous adjustment of the control model with minimal training effort. *iRR-RFF* is effective to retain accurate control across signal shifts [8], [9], and the computational cost of each update is not dependent on the amount of previously trained data, so that model updates remain consistently quick. This is an important efficiency benefit, that makes the method suitable for daily use. Although this mechanism stabilises reliability over time, it still relies

heavily on human-driven feedback. The user will need to permanently monitor performance and intervene for updates. In order to facilitate day-to-day use of an electric prosthesis, it would be desirable to develop a complete system, that assesses its state automatically and triggers updates when needed. In an attempt to automate the incremental model updates, the authors in [10] trained a standard linear classifier to detect failures in the sEMG input signal. The model classified examples of myocontrol use in different tasks as good versus poor control performance, and was able to match a human observer’s assessments with an overall accuracy of 76.71%. This classification relies mostly on detecting features that are generally not desired, such as oscillatory behaviour or high accelerations. It remains unclear whether this classifier could detect a shift in sEMG patterns, that results in plausible predictions, but produces the wrong hand configuration.

The underpinning idea of this work is that automatic failure detection should instead be able to spot every instance, where the myocontrol output does not match the user’s intention. Our hypothesis is that we can detect more subtle shifts in the control mapping by incorporating *situational context information*. In the reach-to-grasp scenario, the hand configuration we choose greatly depends on the shape of the object and how we position our hand relative to the object (i.e. the local context). From human demonstration we learn the relation between context and reach-to-grasp trajectories as a set of generative models. These models are then used as a prior estimate of the user intention in similar local contexts. By comparing new motions to our models’ predictions, we can assess which of the movement commands do not reflect the intention and are more likely caused by a myocontrol failure. We developed and demonstrated our approach in a virtual reality (VR) simulation, where both patients and able-bodied users can control a model of a prosthetic hand. The simulation provides a highly controlled test environment, and allows us to accurately measure, record, evaluate and visualise the user’s movements in the virtual scene.

This paper is structured as follows. We first describe the implementation of the VR simulation. We then show how to construct the proposed approach for learning a situational context model. Finally, we demonstrate that our approach appropriately distinguishes between sets of successful and failed control performance, and can be used to detect myocontrol failures.

*This work was supported by UK Engineering and Physical Sciences Research Council (EPSRC No. EP/R02572X/1).

¹Karoline Heiwolt is with the Computational Neuroscience and Cognitive Robotics (CNCR), University of Birmingham, UK KXH779@student.bham.ac.uk

²Claudio Zito and Rustam Stolkin are with the Extreme Robotics Lab (ERL), University of Birmingham, UK <C.Zito,R.Stolkin>@bham.ac.uk

³Markus Nowak and Claudio Castellini are with the Institute of Robotics and Mechatronics, DLR — German Aerospace Center, Germany <Markus.Nowak,Claudio.Castellini>@dlr.de

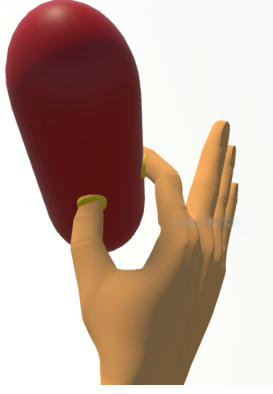


Fig. 1. Example of a stable grasp according to our custom grasp stability rules in the VR simulation. Two contact points are marked in yellow on the object surface. The angle between their contact normals is 90 degrees.

II. METHOD

A. VR Simulation

1) *Materials*: The virtual environment was developed using the Unity 3D game engine and was displayed on the HTC Vive Virtual Reality Headset. One HTC Vive Tracker was used for positional tracking of the user's right forearm. The sEMG data was gathered using the Myoband by Thalmic Labs, a stretchable bracelet fitted with 8 EMG sensors.

2) *Myocontrol Implementation*: We specifically simulate control of transradial prostheses (i.e. below the elbow joint). The virtual prosthesis is modelled as a biological hand with 20 degrees of freedom (DoF), that are grouped into two degrees of control (DoC) for the purpose of this work, namely coupled flexion of the thumb, index finger and middle finger, and coupled flexion of the ring finger and little finger. Using the iRR-RFF as in [7], a model is trained to predict two normalised values, representing the proportional activation of each DoC, from 8-dimensional sEMG patterns. This activation vector is predicted at a frequency of 200 Hz. At every rendered frame of the simulation, we translate the latest available activation vector into 20 joint angles using a mapping matrix. The baseline hand configuration for a zero activation vector corresponds to the resting position (configuration a) in Fig. 4).

3) *Object Interaction*: Physics engines are unstable when multiple contacts act on an object, such as in grasping (e.g. [11]). We thus rely on the physics engine only for collision detection, but allow penetration and specify custom rules for grasp stability. We classify a grasp as stable if there are at least two contacts with a minimum angle of 90 degrees between their normals (see Fig. 1). The object is then simply attached to the wrist's coordinate frame and moves with it, as long as the condition is satisfied. This criterion was experimentally tuned to appear as natural as possible and only fulfills the purpose of enabling basic interaction with a rigid object.

B. Failure Detection

We aim to detect potential failures in the myocontrol by learning a set of generative models that encodes a correct behaviour for attempting a grasp. Our training data is represented as a set of N states recorded along all the demonstrated grasping trajectories:

$$D = \left\{ s_j : s_j \in \mathbb{R}^3 \times \mathbb{H} \times [0, 1]^k \right\}_{j \in [1, N]}, \quad (1)$$

where each state s_j consists of the position $p_j \in \mathbb{R}^3$ of the wrist in space, the quaternion $q_j \in \mathbb{H}$ representing the orientation of the wrist, and a vector $\mathbf{M}_j \in [0, 1]^k$ of activation values for k degrees of control that encode the hand configuration.

The quality of our predictor is evaluated as a normalised weighted mean squared error (MSE) between a new state x and our model's prediction based on the training data D . The normalised weighted MSE can be seen as the expected value of the squared error function of our model. We thus define an error function C as follows:

$$C(x, D) = \frac{1}{\sum_{j=1}^N w_j} \sum_{j=1}^N \text{MSE}(\mathbf{M}_x, \mathbf{M}_j) w_j. \quad (2)$$

The difference in hand configurations is measured by the MSE between two activation vectors:

$$\text{MSE}(\mathbf{M}_x, \mathbf{M}_j) = \frac{\langle (\mathbf{M}_x - \mathbf{M}_j), (\mathbf{M}_x - \mathbf{M}_j) \rangle}{k}. \quad (3)$$

The angled brackets $\langle \cdot, \cdot \rangle$ are used to denote the inner product between two vectors. That way, the MSE compares activations for each DoC and appropriately summarises the total degree of divergence in a single value between 0 and 1, where 0 expresses perfect congruence.

This error is calculated between x and each state s_j in D . But for the overall error of x , we want to compare x only to previously demonstrated states in a similar context, that is states in which the prosthesis was positioned relative to the object in a similar way. Therefore we attach a weight w_j to each training state, that represents its relative importance with respect to x . This weight is a function of both the spatial and angular distance between that training state and the current state x :

$$w_j = w_{\text{position}}(p_x, p_j) \times w_{\text{rotation}}(q_x, q_j) \quad (4a)$$

with

$$w_{\text{position}}(p_x, p_j) = \begin{cases} e^{-\alpha d(p_x, p_j)^2} & \text{if } d(p_x, p_j) \leq \delta \\ 0 & \text{otherwise} \end{cases} \quad (4b)$$

$$w_{\text{rotation}}(q_x, q_j) = \begin{cases} e^{-\beta \theta(q_x, q_j)^2} & \text{if } \theta(q_x, q_j) \leq \phi \\ 0 & \text{otherwise} \end{cases} \quad (4c)$$

where $d(p_x, p_j)$ is calculated as the euclidean distance between the two points, and

$$\theta(q_x, q_j) = \cos^{-1} \left(\frac{\langle q_x, q_j \rangle}{\sqrt{\langle q_x, q_x \rangle} \times \sqrt{\langle q_j, q_j \rangle}} \right). \quad (4d)$$

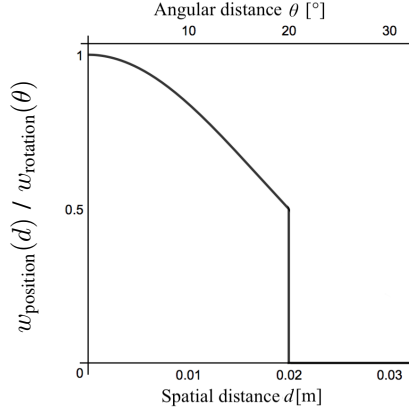


Fig. 2. Weight functions w_{position} with respect to the lower x-axis and w_{rotation} with respect to the upper x-axis, plotted for the parameter values $r = 0.25$, $\delta = 0.02\text{m}$, $\phi = 20^\circ$, $\alpha = 1733$ and $\beta = 0.001733$. Both functions implement the same decrease of relative importance for points within their respective cut-off distances δ and ϕ . In this example the values range between 0.5 and 1, so that the combined weight w_j can assume values between 0.25 and 1.

These two weighting terms decrease exponentially with the square distance in space and in angle respectively. Additionally, training states further than the cut-off distances δ and ϕ from x are excluded entirely. Thus, these two parameters define a focal area (or area of similar context) around x , in which training states are considered to be relevant. Parameters α and β can be tuned to specify how much points on the edge of the focal area contribute to the evaluation relative to points in the centre:

$$\alpha = -\frac{\ln(\sqrt{r})}{(\delta)^2} \quad (5a)$$

$$\beta = -\frac{\ln(\sqrt{r})}{(\phi)^2} \quad (5b)$$

where r is the relative contribution ratio or relative importance of training states on the edge of the focal area (for example see Fig. 2).

Then, we integrate all those single state comparisons into one overall error function C for the state x . In Eq. (2) this is computed as the normalised weighted mean of the differences in hand configurations, weighted by their similarity in context, rather than a sum or an arithmetic mean over the comparisons. The weighted mean appropriately expresses the overall divergence of the hand configuration from previous demonstrations in a similar context. It is not distorted by the amount of available data and takes into account the relative importance of training states.

In summary, the error function given in Eq. (2) can be used to calculate a mismatch between the user's motion and the model's prediction in a similar context. The learned models can be interpreted as an estimate of user intention, thus the error values also encode how much the performed action differs from the command we assume the user might have attempted. We can also infer whether the activation vector in the new state is likely an accurate reflection of the user's intention (low values) or rather a failure in the myocontrol

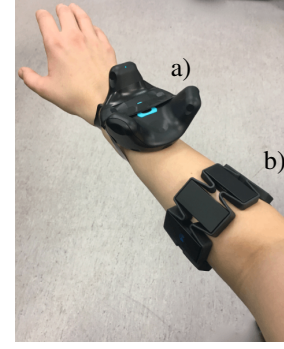


Fig. 3. Placement of the sensors on the user's arm: a) The Vive Tracker is placed just below the wrist joint and secured using adhesive tape, b) The Myoband is placed approximately 5cm below the elbow joint, where it picks up electrical signals at 8 positions around the forearm. Exact placement of the electrodes is not necessary for regression model approaches, as long as they cover different muscles around the forearm (e.g. [7], [12], [13]).

(high values). To do so, a threshold can be defined, above which states are classified as myocontrol errors.

C. Experimental Method

The ability of the proposed function in Eq. (2) to detect failures was analysed on a reach-to-grasp task. One set of training data and two sets of test data (a success condition and a failure condition) were collected in a single session by the same able-bodied subject, using the VR simulation described in section II-A.

1) *Procedure:* The sEMG bracelet and the Vive tracker were placed on the user's forearm as shown in Fig. 3.

A myocontrol model was trained on the three different hand configurations shown in Fig. 4. Just as in [7] and [13], the subject performed the desired grasps by copying the visual cues in Fig. 4 at the maximal comfortable level of force. The resulting sEMG activation pattern was labeled with the two-dimensional activation array corresponding to the displayed grasp. In order to achieve an accurate reflection of the user's intention, the model was carefully adapted incrementally until the subject was satisfied with the control performance and subjectively perceived no control failures. In total, each grasp was recorded eight times, four times in a relaxed position with the right arm held close to the body and bent by 90 degrees at the elbow, and four times with the arm extended forward and the palm facing inward. Once the system was trained, the subject put on the head-mounted display and started the data collection.

For the training data, the subject performed a total of 22 trials of the grasping task. In each trial a capsule-shaped object appeared on a table and within reaching distance of the user in the virtual scene. The subject then reached for the object and picked it up, while the application recorded the trajectory and configuration of the virtual hand continuously until the first stable grasp on the object was achieved. The capsule shape was chosen, as it resembles many real life demands (such as picking up a bottle), which can be solved in a number of different ways, including at least two different grasp types (power grasp and tridigital grasp)

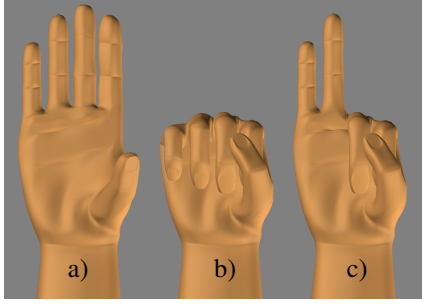


Fig. 4. Hand configurations used for training in the experiment: a) Resting position b) Power grasp c) Tridigital grasp.

and several different contact regions and arm configurations. So instead of repeatedly performing the same motion, the subject was encouraged to demonstrate a variety of different grasp solutions on the object. For a visualisation of the training data set, see the left panel of Fig. 5.

Because this training data serves as a sample of natural grasps a human user would attempt, it is essential to avoid accidentally recording any myocontrol failures within this sample. Thus, all indistinct trials, that is trials in which the myocontrol did not exactly reproduce the subject’s intention, were discarded immediately, per the judgment of the subject. Twice during this data collection a single update was incrementally added to the myocontrol model to counteract performance degradation.

Next, we collected test data for the validation of the proposed error function. The success condition is a sample of successfully executed grasps, representing the desired state of a reliable myocontrol system, and containing no myocontrol failures. For this sample 8 trials were collected in exactly the same way as described above, where the execution accurately reflects the subject’s intention. For the failure condition, we aimed to collect data, that represents the myocontrol failures in the unreliable state of the myocontrol system following a shift in the sEMG inputs. A second myocontrol model was trained exactly as before, but this time the labels (i.e. the activation vectors) for the grasps b) and c) in Fig. 4 were swapped. By intentionally mis-labelling the grasps, we effectively trained an action to be associated with a different intention. The resulting activations and model predictions were plausible, but produced a different hand configuration from what the user intended. Another 20 trials were recorded with this second myocontrol model.

2) *Data Analysis:* The three data samples were represented as arrays of states (as defined in Eq. (1)), sampled at approximately every 12 milliseconds along each trajectory. We then conducted two evaluations of both experimental conditions using the function specified in Eq. (2). For the first evaluation, we calculated the value of the function for every single test state in both experimental conditions based on the full training data; a total of 2789 training states, shown in the left panel of Fig 5.

The second time, we calculated the errors based on a subset of the training data, containing only the last 20

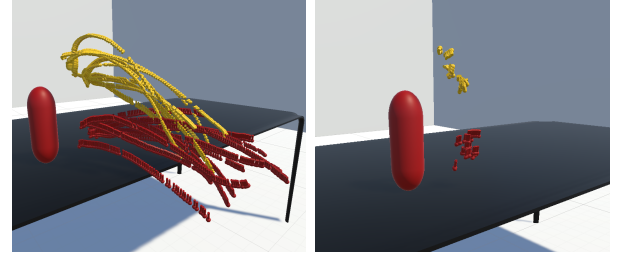


Fig. 5. Visualisation of the training data. The left panel shows 2789 states sampled along the whole length of all demonstrated trajectories. The right panel shows the 440 target area training states, representing only the last 200ms of each trajectory. Trajectories pictured in yellow ended in a tridigital grasp, red trajectories ended in a power grasp.

TABLE I
PARAMETER CHOICES FOR THE ERROR FUNCTION

k	n	r	δ	ϕ	α	β
2	5	0.25	0.02m	20°	1733	0.001733

TABLE II
OVERVIEW OF TEST SAMPLE SIZES BASED ON BOTH POSITIONAL AND ORIENTATIONAL INFORMATION OF THE WRIST.

	Success condition	Failure condition
Number of grasp trials	8	20
Total number of sampled test states	1099	2066
Number of evaluable test states based on the full training data	605	765
Number of evaluable test states based on the target area training data	306	338

recorded data points before the stable grasp was reached in each demonstrated trial. This smaller *target area training set* contains a total of 440 training states representing roughly the last 200ms of each trajectory, as shown in the right panel of Fig 5. This second evaluation was conducted, because in qualitative observations, most trajectories in both conditions started in a very similar way and only diverged into different hand configurations close to the object. We expected that the difference in error distributions between the conditions also becomes more pronounced closer to the object surface. Thus, evaluating the test data only on the smaller but more determining and relevant target area training sample should suffice to distinguish between the two conditions. Finally, all test states for which there were not enough training states n available in a similar context for an evaluation ($n < 5$) were excluded from further analyses. All function parameters were tuned experimentally and kept fixed throughout these experiments (Table I). For an overview of the final sample sizes, see Table II.

We then compared the sets of calculated error values across the two conditions. Overall differences in the distribution of errors were analysed using statistical tests. Furthermore, we closely examined the absolute range of the values, as this feature in particular determines how appropriate a threshold would be as a final failure detection criterion. One possible concern of this approach is, that single trials

TABLE III

COMPARISON OF ERROR DISTRIBUTIONS BASED ON BOTH POSITIONAL AND ORIENTATIONAL INFORMATION OF THE WRIST.

Evaluation 1	Success condition	Failure condition
M	0.0129	0.0437
SD	0.0269	0.1066
Median	0.0054	0.0120
p	<0.001	
Evaluation 2	Success condition	Failure condition
M	0.0247	0.0948
SD	0.0317	0.1489
Median	0.0104	0.0391
p	< 0.001	

Note. Values reported in this table are the mean (M), standard deviation (SD), and the median of all calculated errors in each experimental condition, as well as the one-sided p -value of a permutation test between the conditions after 5000 permutations.

could be over-represented within their samples. Because of different movement speeds during the data collection, the number of states within single trials varies, and we need to ensure that the overall results in the direct comparison are not distorted, but that similar results are obtained when the data is grouped into trials for analysis. Therefore, we also examined the overall distributions, the mean error and the ranges along each trial. It is important to note however, that the error is intended to be interpreted with respect to the single current state, and this grouped analysis only serves as a validation of direct comparison results. And finally, we tuned an exemplary threshold value to the results and determined the sensitivity and specificity of the eventual failure classification of our test data set.

To translate this approach into a mobile prosthesis, the device needs to be fitted with a depth camera and on-board positional and rotational tracking that does not require external reference (for example inertial tracking, e.g. [14]). Since this positional information is much harder to obtain than orientational information, we also assessed the distribution of errors and the potential failure detection performance based only on the orientation of the wrist, in case there is no positional information of the wrist available at all.

III. RESULTS

A. Based on both positional and rotational information

The distributions of the errors for both conditions in both evaluations are shown in the left two groups of Fig. 6. Most of the data accumulates around the low end of the spectrum. Overall, however, the failure condition yielded higher values than the success condition (Table III). This difference was further analysed using a permutation test. Specifically, we used a MatLab implementation by Ehninger [15], based on the corrections proposed by Phipson and Smyth [16]. The test detected statistically highly significant differences in error distributions with $p < 0.001$ after 5000 permutations for both evaluations.

We then examined the same data grouped into trials, and the effect persists. The mean error values, when corrected

TABLE IV

SAMPLE SIZES BASED ON ORIENTATIONAL INFORMATION OF THE WRIST ONLY.

	Success condition	Failure condition
Number of grasp trials	8	20
Total number of sampled test states	1099	2066
Number of evaluable test states based on the full demonstration data	1099	2066
Number of evaluable test states based on the target area demonstration data	1099	2051

for the different amounts of evaluable states in single trials, describe the same relation between the two conditions. The mean range along the trials also differed between the conditions just as expected from the overall comparison between the groups on both evaluations (Fig. 7). After a closer look at the descriptive data for each trial, we conclude that there is no reason to believe that the difference found in the overall comparison misrepresents the underlying data.

The range in error values in the failure condition is substantially wider than in the success condition by a factor of 2.65 in the first evaluation, and 5.15 in the second evaluation. Fig. 6 also shows that the difference between conditions is more pronounced in the second evaluation (based on only target area training data) than in the first evaluation (based on all training data). Accordingly, we observed that the error development along single trials also shows a greater divergence between the two conditions towards the end of each reach-to-grasp trajectory (Fig. 8).

We also found that the maximal error value along each trial is reached very close to the object surface. In the first evaluation, across all grasp trials the highest error value is reached on average only 9.75 states away from the final grasp in the success condition, and 3.15 states in the failure condition. Similarly, in the second evaluation, the highest errors are on average 14.63 (success) and 4.85 (failure) states away from the grasp. All of these averages lie within the last 176ms or less of each trial. Additionally, we found that the reported differences and especially the difference in range can be further tuned, by varying other parameters of the error function. As an example, the third group in Fig. 6 displays error distributions for an evaluation using the target area training data and $n = 20$.

Using a threshold of $C_T = 0.0249$ on our test data (evaluated on the target area training data with $n = 5$), the method correctly detected 18 out of 20 failures trials and correctly classified 6 out of 8 correct trials as non-failures (Sensitivity = 0.9, Specificity = 0.75).

B. Based on rotational information only

In this second analysis the focal area of similar context was determined only by the angular distance between a test state and the training data. As a result, more training states were considered to be relevant to the current situational context. This greater generalisation increased the number of evaluable states (Table IV), but at the same time slightly reduced the

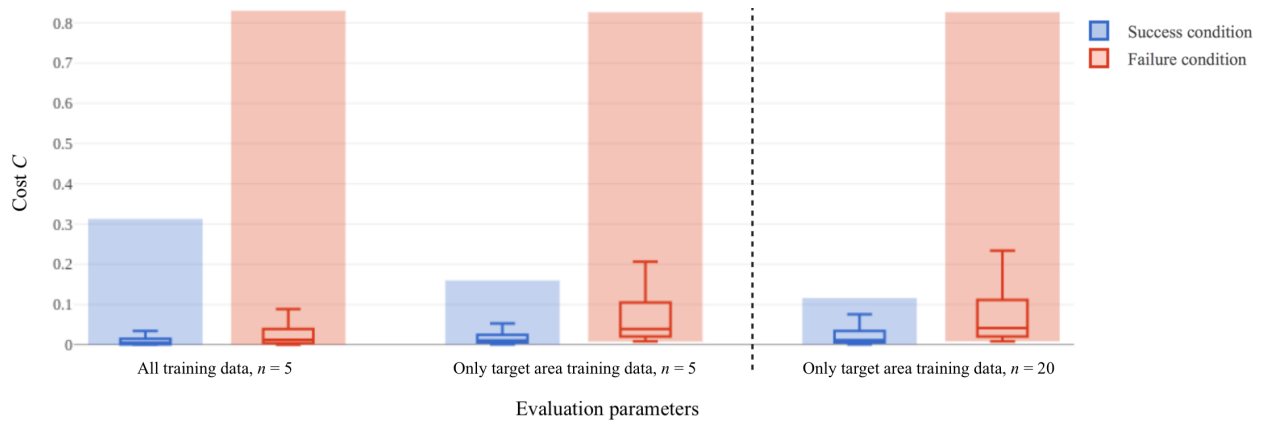


Fig. 6. Comparison of error distributions between the two experimental conditions when both positional and rotational information were available. The chart shows Tukey boxplots, where the whiskers represent data within 1.5 inter-quartile-ranges of the first and third quartile. The bars illustrate the absolute range of error values. The two groups to the left display the results obtained in evaluation 1 and 2, while the group on the right represents just one example of how error distributions can be influenced by tuning other variables of the error function.

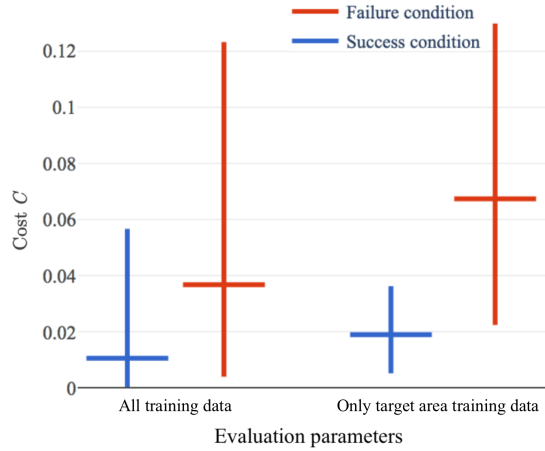


Fig. 7. Results of the error evaluation for all test states grouped into trials. The horizontal lines indicate the mean value, corrected for different numbers of evaluable test states across single trials. The vertical lines indicate the average range along single trials in the condition.

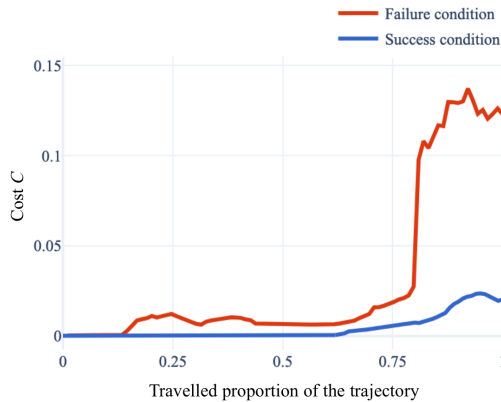


Fig. 8. Error values along one single exemplary grasp trajectory of each experimental condition.

divergence between the two experimental conditions (Table V, Fig. 9). However, there is still a clear distinction between the success and the failure condition. Using a threshold

TABLE V
COMPARISON OF ERROR DISTRIBUTIONS BASED ON ORIENTATIONAL INFORMATION OF THE WRIST ONLY.

Evaluation 1	Success condition	Failure condition
M	0.0192	0.0436
SD	0.0231	0.1022
Median	0.0189	0.0217
p	<0.001	
Evaluation 2	Success condition	Failure condition
M	0.0533	0.0719
SD	0.0307	0.0802
Median	0.0454	0.0546
p	< 0.001	

Note. Values reported in this table are the mean (M), standard deviation (SD), and the median of all calculated errors in each experimental condition, as well as the one-sided p -value of a permutation test between the conditions after 5000 permutations.

value of $C_T = 0.0511$ on our test data (evaluated on all training data) the method achieved a sensitivity of 0.6, and a specificity of 0.875.

IV. DISCUSSION

The experimental results indicate that the proposed approach can potentially be used to monitor the status of a prosthetic control system, detect and even anticipate failures, and therefore increase its accuracy and reliability over time. The model can clearly distinguish between the two conditions; while lower values occur in both conditions, for high values it becomes more and more likely that the observed state is a failure, as there is a large range of error values that only ever occurred in the failure condition (see Fig. 6). The way to achieve an optimal failure detection must therefore be to minimise the overlap of the two conditions, or the area of uncertainty, and tune the error calculation in such a way, that the data in the two groups diverge as much as possible.

One simple way of determining whether we are facing a failure is then to use a threshold for triggering a model update; such a threshold will always constitute a trade-off

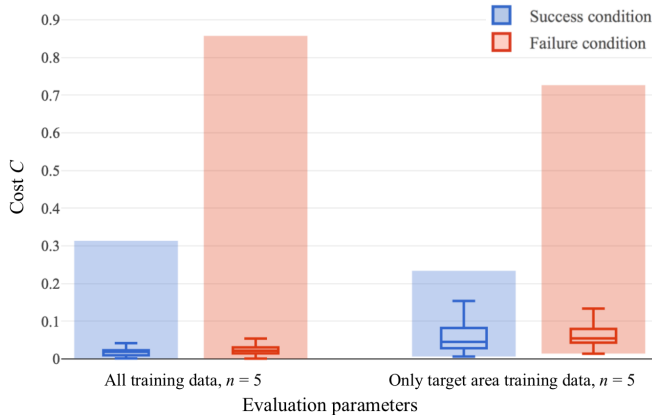


Fig. 9. Comparison of error distributions between the two experimental conditions, based on only orientational information of the wrist. The chart shows Tukey boxplots, where the whiskers represent data within 1.5 inter-quartile-ranges of the first and third quartile. The bars illustrate the absolute range of error values.

between sensitivity and specificity, but the more distinctly we can isolate the two conditions, the more accurately and confidently we can place the threshold and identify failures. The exact value could even be tuned individually, depending on the available data and user preference.

In this work we have made several assumptions. Firstly, the experiment was limited to two grasp types only, and secondly, the task only included a single object shape. Additionally, the approach in this paper was tested only on one non-disabled expert user. Future work will further explore the proposed method in a broader experiment with more participants, including patients from the target population. We also assumed that the user attempts a sensible grasping motion, as opposed to e.g. punching the object. In such cases the system will detect a failure, but the user should be able to dismiss the alert.

As future extension we will also build on previous work of Kopicki and Zito [17], [18], [19], [20] who have demonstrated that a set of generative models can be efficiently learned, for a robot manipulator, in one shot such that manipulative contacts and trajectories are computed for previously unseen objects. By integrating these methods in our system, we will replace the human demonstration samples as a prior estimate of which grasps the user might attempt, and more importantly it would allow generalisation to new object shapes.

V. CONCLUSION

In this work we introduced an automatic failure detection method that continuously monitors the system performance. Although we are still in an early stage of development, our results indicate that the approach performs well and has potential for future extensions. We believe that this is a promising first step towards a truly reliable self-correcting control system, highly interactive and customisable, without the need for excessive training.

REFERENCES

- [1] B. Peerdeman, D. Boere, H. Witteveen, R. H. in 't Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, and S. Misra, "Myoelectric forearm prostheses: State of the art from a user-centered perspective," *The Journal of Rehabilitation Research and Development*, vol. 48, no. 6, p. 719, 2011.
- [2] S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 48–68, 2010.
- [3] A. Fougner, O. Stavdahl, P. J. Kyberd, Y. G. Losier, and P. A. Parker, "Control of upper limb prostheses: Terminology and proportional myoelectric control—a review," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 663–677, 2012.
- [4] N. Jiang, S. Dosen, K.-R. Muller, and D. Farina, "Myoelectric control of artificial limbs—is there a need to change focus? [in the spotlight]," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 152–150, 2012.
- [5] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014.
- [6] J. M. Hahne, M. A. Schweisfurth, M. Koppe, and D. Farina, "Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users," *Science Robotics*, vol. 3, no. 19, 2018.
- [7] A. Gijssberts, R. Bohra, D. Sierra González, A. Werner, M. Nowak, B. Caputo, M. A. Roa, and C. Castellini, "Stable myoelectric control of a hand prosthesis using non-linear incremental learning," *Frontiers in Neuroinformatics*, vol. 8, 2014.
- [8] D. Sierra González and C. Castellini, "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees," *Frontiers in Neuroinformatics*, vol. 7, 2013.
- [9] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 788–797, 2012.
- [10] M. Nowak, S. Engel, and C. Castellini, "A preliminary study towards automatic detection of failures in myoelectric control," in *MEC17 - A Sense of What's to Come*, 2017.
- [11] K. Nasim and Y. J. Kim, "Physics-based interactive virtual grasping," in *HCI Korea 2016*. The HCI Society of Korea, 2016.
- [12] C. Castellini, E. Gruppioni, A. Davalli, and G. Sandini, "Fine detection of grasp force and posture by amputees via surface electromyography," *Journal of Physiology*, vol. 103, no. 3-5, pp. 255–262, 2009.
- [13] I. Strazzulla, M. Nowak, M. Controzzi, C. Cipriani, and C. Castellini, "Online bimanual manipulation using surface electromyography and incremental learning," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 227–234, mar 2017.
- [14] X. Yun, E. R. Bachmann, H. Moore, and J. Calusdian, "Self-contained position tracking of human movement using small inertial/magnetic sensor modules," in *Proceedings IEEE Int'l Conf. on Robotics and Automation*, 2007.
- [15] B. Ehinger. (2016) Permtest. [Online]. Available: <https://github.com/behinger/permtest>
- [16] B. Phipson and G. K. Smyth, "Permutation p -values should never be zero: Calculating exact p -values when permutations are randomly drawn," *Statistical Applications in Genetics and Molecular Biology*, vol. 9, no. 1, 2010.
- [17] M. Kopicki, R. Detry, M. Adjigble, R. Stolkin, A. Leonardis, and J. L. Wyatt, "One-shot learning and generation of dexterous grasps for novel objects," *The Int'l Journal of Robotics Research*, vol. 35, no. 8, pp. 959–976, sep 2015.
- [18] J. Stuber, M. Kopicki, and C. Zito, "Feature-based transfer learning for robotic push manipulation," in *Proceeding of IEEE International Conference on Robotics and Automation (ICRA)*, 2018.
- [19] C. Zito, M. Kopicki, R. Stolkin, C. Borst, F. Schmidt, M. A. Roa, and J. L. Wyatt, "Sequential trajectory re-planning with tactile information gain for dexterous grasping under object-pose uncertainty," in *Proceeding of IEEE International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 2013–2040.
- [20] C. Zito, V. Ortenzi, M. Adjigble, M. S. Kopicki, R. Stolkin, and J. L. Wyatt, "Hypothesis-based belief planning for dexterous grasping," *arXiv preprint, arXiv:1903.05517 [cs.RO] (cs.AI)*, 2019.